CONTINGENT VALUATION. A Critical Assessment
J.A. Hausman (Editor)

Chapter I

ON CONTINGENT VALUATION MEASUREMENT OF NONUSE VALUES

PETER A. DIAMOND and JERRY A. HAUSMAN
Massachusetts Institute of Technology

INTRODUCTION[1]

Americans are concerned about the effects of the economy on the environment. Individually, some of us show this concern by donating to environmental groups. Collectively, we show this concern through government regulations and government expenditures designed to protect the environment. To help guide these regulations and expenditures in the direction of greater protection, some economists have developed the concept of "nonuse value." Although we shall define nonuse values, we first give some possible examples of nonuse values:

(1) The value of the Grand Canyon to you if you never intend to go to the Grand Canyon.
(2) The value of a "pristine wilderness" in northern Maine that may visited by few, if any, people.
(3) The value of a clean lake in extreme northern Ontario.
(4) The value of a nearly extinct species, such as the spotted owl.
(5) The value of an otter or a dolphin, members of species that are not endangered.

As these examples illustrate, nonuse values do not influence individual actions the same way that values for economic goods influence the purchase and sale of such goods. Economics as an intellectual discipline makes actions by individuals and firms the primary focus of study. Thus, individual actions, such as the recreational use of a given environmental site, demonstrate the value individuals place on the site. Economists can estimate this value by using the technique of revealed preference,[2] whereby preferences are inferred from behavior. However, all of the examples of nonuse values given above are difficult to measure because the commodities being valued are not traded in markets, nor are individual actions affected by the particular

nonuse values. Thus, even if some people have preferences that imply nonuse values, it is difficult to put a price tag on something that is never traded and does not affect individual actions in the normal manner.

The measurement of nonuse values is potentially important because economic analysis and the courts use similar values in reaching decisions. For instance, benefit-cost analysis is a method developed by economists to provide information to be used in government decision making. Benefit-cost analysis is applied to determine rules for spending on public goods (e.g., Samuelson, 1954) or to set regulations, for example, on permissible levels of pollutants. Courts use measurements of values to set compensatory damages for accidents. Rules used by the courts to measure tort damages can encourage economic efficiency (Shavell, 1987). However, concerns about the environment that may be reflected by nonuse values have not been addressed fully by traditional economic analysis in benefit-cost analysis or in assessing compensatory damages.

Inclusion of nonuse values is certainly not required in these forms of economic analysis. Individuals have concerns about the poor, similar to concerns they hold for the environment. Furthermore, governments make expenditures on behalf of both the environment and the poor. Yet, neither benefit-cost analysis nor damage assessment in the courts has traditionally incorporated concerns about income distribution or the nonuse values of the environment.

Recently, an attempt to incorporate nonuse values into benefit-cost analysis and damage assessment has begun. Measurement techniques for nonuse value have relied on the contingent valuation (CV) method. In this paper we briefly review the concept of nonuse value and then assess the attempt to measure nonuse value by the CV method. We examine evidence on the lack of consistency of these CV measurements with the economic theory on which the measurements are based and also include reports on some new CV experiments.

These new experiments were meant to focus specifically on the consistency of CV answers with basic economic principles. Thus, the main focus of the paper is the question of whether nonuse values measured by CV methods are internally consistent with the economic theory on which they are based. We conclude that the CV method does not measure an economic value that conforms with economic preference concepts. Thus, we also conclude that it is not appropriate to include CV measures of stated willingness to pay (WTP) in either benefit-cost analysis or compensatory damage measurement. We go on to venture a possible framework for interpreting the responses to CV questions.

The importance of reviewing the CV approach to measuring attitudes to the environment arises, in part, because of its incorporation in the regulations of the Department of the Interior (U.S. Department of the Interior, 1986) for measuring the damages associated with oil spills and hazardous wastes under the Clean Water Act of 1977 (CWA), the Comprehensive Environmental Response, Compensation and Liability Act of 1980 (CERCLA), the Superfund Amendments and Reauthorization Act of 1986 (SARA), and the Oil Pollution Act of 1990 (OPA).

Therefore, we focus on the question of whether legislation that calls on the Department of the Interior to measure the damages needed to compensate people for natural resource injury ought to include CV measurement of nonuse values. The stakes in this issue are large because usual CV measurement of nonuse values for an area such as Prince William Sound are likely to exceed all conventionally measured damages that are based on revealed preference actions by individuals (e.g., lost fishing and use values from recreation) by one to several orders of magnitude.[3] Of course, our view that it is not possible to generate a meaningful estimate of nonuse values suitable for benefit-cost analysis or for compensatory damages does not preclude legislation to protect the environment that incorporates financial incentives such as fines, as well as the more familiar direct controls. We conclude that fines and direct controls are a better public-policy response than government actions that are based essentially on a method that relies on the results of opinion polls over unfamiliar choices, which is essentially what we interpret contingent valuation to be.

HISTORY AND DEFINITION OF NONUSE VALUES[4]

History

The intellectual origin of nonuse values lies in the article by Krutilla (1967).[5] After raising concern about the quality of the physical environment, Krutilla began his analysis: "Let us consider an area with some unique attribute of nature—a geomorphologic feature such as the Grand Canyon, a threatened species, or an entire ecosystem or biotic community essential to the survival of the threatened species" (p. 778).[6]

In contrast with this beginning, some economists have argued that nonuse values are applicable to everything in sight. For example, Randall and Stoll (1983) have argued:

> Thus, even commonplace artifacts of human civilization (e.g., drink cans) may have existence value, although the circumstances which would make it large are unlikely. Empirically significant existence values are not confined to natural objects; we believe they occur for human artifacts and cultural manifestations, from historic buildings to grand opera. (p. 268)

Several things should be noticed about the perspective of Krutilla's beginning. One is that the focus is on unique (unusual) and collectively important sites, such as the Grand Canyon. This focus also suggests that it is the long-run or permanent consequences for a site that are of primary importance, not temporary harm. One possible exception to the distinction between temporary and permanent (or long-run damage) comes when a specific value is placed on the fact that a particular site has never been damaged. That is, individuals may associate a value with the pristinity

of a site. Such concern will often not survive the reality that pristinity is nonexistent for almost all U.S. sites.

Use versus Nonuse Values

A large body of economic literature has been developed on the treatment of marketed goods in both benefit-cost analysis and damage measurement. This literature is based on the standard presumption of economists that more is to be learned from observations on how people behave rather than from answers to hypothetical questions. Thus, watching the behavior of people in market situations is the heart of traditional benefit-cost analysis, and, for that matter, most of economics. When someone is observed purchasing a good at some price, it is inferred that the good is worth at least its cost to the purchaser.[7] If the good is then destroyed in an accident, the price that has been paid for it is a natural measure for the damages inflicted, especially if the person who caused the accident is capable of replacing the good at the same price. For nonmarket goods, such as natural resource sites, the analysis often becomes more difficult because they have no market price. But, straightforward damage measurement can still be undertaken if it is possible to link underlying preferences with observed behavior, as in market goods.

Thus, evaluation of the use value of a site, for example, for recreational fishing, parallels conventional purchases of marketed goods. The complexity that arises for analytical purposes is the lack of market prices for the use of such resources. Thus, one cannot directly infer preferences from observed purchases at observed prices. To measure use values, the literature responds to the absence of market prices by making inferences from the costs that do exist for usage of alternative sites. These techniques are known as travel-cost methods. However, the basic framework of the travel-cost method continues to rely on observable actions by individuals—by observing the costs associated with choosing one site over alternatives, economists infer the preferences over the different sites.[8] For instance, after an oil spill, individuals may shift their recreational fishing activity to a site not affected by the spill or decrease the number of recreational fishing trips made.[9] In this estimation of use values, individuals reveal their preferences by their actions, so that the only unusual feature of the situation is that no explicit price is paid for the use of the resource.

Nonuse value is meant to complement use value in considering the total value of preserving a site. Nonuse values have been described in terms of a variety of individual concerns, a number of which will be described below. Although many ways of categorizing different nonuse values have been advanced in the literature, we shall categorize them by considering: (1) values held by an individual related to the individual's own use; (2) values held by an individual related to the use by others; and (3) values held by individuals that are not connected to human use. Two questions need to be asked about any particular nonuse value. First, can the particular nonuse value be measured with reasonable accuracy? Second, is the particular nonuse value a suitable candidate for inclusion in the economic theories

of benefit-cost analysis or damage measurement (for either compensation or incentive purposes)?

Nonuse Values from an Individual's Own Use of a Site

We first consider the values held by an individual related to the individual's own future use of a particular site. The historical and projected use of a resource is the basis for estimating use value. However, calculating the expected value of the number of future uses may not be a precisely correct basis for benefit-cost analysis when individuals are uncertain about their plans with respect to their future use of a site.[10] Option value is the amount that individuals would pay to preserve the opportunity of using the site in the future, in addition to the expected value to them of their uses of the site.[11] Thus option value is a correction term added to expected value calculations. This correction is inherently complex and may be either positive or negative.[12] We expect that, in most situations, this correction will not be large relative to the expected value of future use of the site because it is based on the expected value of future use.[13]

A second element in individual decision making arises from the fact that one may learn over time about the desired use of a facility. For example, if a dam is built that floods an unexplored canyon, the opportunity to discover possible important archaeological ruins is lost. Thus, the value of possible additional information changes the correct benefit-cost calculation. This concept differs from the option value described above because of the explicit recognition of the link between policy actions and possible additional learning. The value associated with this correction to the expected value calculation has been called the quasi-option value.[14] Although learning occurs in many circumstances, quasi-option value has been thought to be particularly important in the circumstance of an irreversible change in a site. The theoretical literature[15] demonstrates that quasi-option value can be positive or negative, because different patterns of learning occur with different uses of a site. Quasi-option value is unlikely to be important for temporary damage, as typically occurs with an oil spill, because often no irreversible changes occur.

Nonuse Values Related to the Enjoyment of a Site by Others

In considering the values individuals hold for a site to be used by others, it is natural to begin with bequest value.[16] Bequest value, when restricted to, for example, a couple's descendants, is simply the recognition that benefits conferred by the future use of a resource extend beyond their own enjoyment to the enjoyment of their children.[17] Because the economic literature has not succeeded in separating the relative empirical importance of unintentional bequests from planned bequests, except among the very wealthy, it is unclear how large this element is likely to

be. When one calculates the present discounted value of the future use of a given site for the rest of time, which is done with estimates of the current and future use values of a given site, one would automatically be counting the use by future generations. It would appear to be double counting to count both the future generation's use and the current generation's pleasure from the future generation's use.[18]

Whenever a resource is used by more than one person, the value to each person of using the site may increase or decrease as a result of its use by others. Economists call this change in value because of the use by others an "externality."[19] Use-related externalities are naturally part of the evaluation of the enjoyment of the site, and they will be measured with traditional methods. Other externalities parallel the bequest value, reflecting interest in the use by others, without direct interaction. Externalities thus involve the same sort of double counting as was described above for the use-related part of the bequest value. Such concern for others is normally excluded from both benefit-cost analysis and from damage measurement.

All of the values described above relate either to the direct or indirect use of a resource by a given individual. Paying attention to these additional indirect use elements is unlikely to change significantly the measurement of resource values beyond conventional usage-based measures if neither you nor your descendants are likely to use a remote site. For damage assessment, this conclusion is stronger when the damage is temporary, so that the change in future or potential uses will not be large. The central question at hand is if nonuse value unrelated to the use of others has an independent, large role.

Values Unrelated to Human Use of a Site

Individuals do give donations to charitable causes to protect resources they have little or no expectation of using. Individuals also support government actions to set aside wilderness areas they never expect to visit. Thus a presumption exists that there are values held by individuals beyond the values related solely to the use of natural resources. Our focus in this paper is on such nonuse values, rather than on examining in any detail the other elements discussed above. In practice, recent CV studies have asked about total valuation of a resource, without attempting to separate out different components.[20] The CV studies reported in this volume also ask questions only about total valuation. Because total value responses are frequently much larger than separately measured use values, the total valuation answers must be predominantly nonuse values. In addition, we see little reason to suggest that the nonuse values associated with quasi-option value or the use by descendants would be large when use value is small; thus, we infer that the nonuse value unrelated to human use is the critical one.

How Can Nonuse Values Be Integrated into Economic Theory?

Let us consider preferences as they are commonly described in economics. When thinking about conventional goods (cars, televisions, clothing), we can easily imagine preferring one set of goods over another. Now add to these conventional goods preferences over natural resources—preferences that include nonuse concerns.[21] Also, initially, we assume that these individual preferences satisfy the normal axioms of well-behaved preferences in basic economic theory, whereby preferences are assumed to be continuous and transitive.[22]

Economic models of consumer behavior describe how individuals make purchases of goods according to their preferences and their income (or budget) constraint.[23] From this perspective, the state of the environment affects well-being but is not a commodity over which an individual exercises choice. Whether changes in the state of the environment do or do not affect demands for ordinary commodities is not relevant for CV studies.[24]

We recognize that individuals have opportunities to donate to charitable causes, that is, to give up income usable for the purchase of ordinary commodities in order to contribute to an improvement in the condition of environmental resources. We return to this behavior below when we interpret some of the empirical findings.

Economists have attempted to deal with the problem that individuals have ethical views in addition to self-interested preferences. A commitment to these ethical views can influence individuals' behavior (Sen, 1977). For instance, in considering government action to affect the distribution of income, individuals may advocate a government action that goes against their self-interest because it coincides with their ethical views. Likewise, ethical views can be quite important for understanding how people answer CV questions. Respondents to CV questions may include an ethical component in addition to a component derived from self-interested preferences.[25] But, benefit-cost analyses and compensatory damage assessments should not take into account ethical values; instead, they should be based only on self-interested (economic) preferences. Monetary compensation for natural resource injury, for instance, should be based only on losses to individuals derived from their preferences for the resource, not on ethical values concerning the injury. Thus, to the extent that CV answers include an ethical component, CV surveys are unsuitable for assessing the appropriate level of compensatory damages.

POSSIBLE PROBLEMS WITH THE CONCEPT OF NONUSE VALUES

Information and Level of Well-Being

When preferences are defined solely in terms of the commodities that people

purchase and use, those commodities unknown to the individual play no role in the individual's level of well-being. Advertising serves, in part, to make individuals aware of the existence and characteristics of additional commodities, enabling them to achieve a higher level of well-being, given their budget constraints. Individual consumers devote effort to acquiring information about commodities and about opportunities to purchase them. Economic theory often concentrates on the individual decision process of how much of each good to purchase, assuming perfect information about available commodities. However, a more realistic model describes an individual with a limited level of information about consumer opportunities, who decides both how much to spend in learning about consumption opportunities and how to pursue the process of seeking out and buying the desired goods. Thus, information acquisition forms part of the overall individual consumption process.

Information plays a different role with regard to nonuse values than it does with ordinary commodities. For example, before the *Exxon Valdez* spill, we suspect that relatively few Americans had ever heard of Prince William Sound by name. Although many Americans knew that Alaska contains much wild and beautiful scenery, most individuals probably had very little awareness of the exact nature of the environment affected by the *Valdez* oil spill. This lack of previous knowledge does not imply that it would be inappropriate to recognize the concern of individuals for the environment. However, the knowledge issue creates a clear difficulty in attempting to measure compensatory damages for the loss of nonuse value when an individual learns simultaneously about a resource's existence and about an injury to it, for example, learning about the existence of Prince William Sound *and* its beauty and the oil spill in the same news report. The change in well-being when a known resource is injured is not the same as that which occurs when one learns simultaneously about the existence of a resource and an injury to it. Is an individual worse off with these two pieces of knowledge than with no knowledge at all?

Information on Outcomes Rather than Inputs

CV studies assume that preferences are defined over resources and their condition, but they do not assume that preferences are defined over the different ways in which a resource might be injured. Whether a beach is oiled by seepage from the seabed or from a spill from a passing oil tanker should not alter individual valuation of the difference between a nonoiled beach and the same beach after oiling. Furthermore, by defining preferences over resources, if a resource is only temporarily damaged, the utility it provides returns to the same level after the damage is eliminated.

Incorrect Information and Knowledge Formation

Further difficulties in trying to evaluate the importance to individuals of

environmental injuries arise from the limited and/or possibly incorrect knowledge that people have about the effects of particular events (such as oil spills) that cause such injuries. Presumably, before determining public policy, we would want to correctly understand the effects of the events.[26]

Advertising affects consumers' knowledge as well as their behavior, so we can imagine the difference between demand for a product before advertising and demand after advertising. But in damage measurements, we take the individuals' positions, including awareness, as given, rather than hypothesizing how they might behave if they had different levels of knowledge. In the case of an oil spill, this general approach suggests focusing on the level of knowledge of individuals before the spill rather than after widespread television coverage of the event. It is difficult to see how one could assess the value individuals place on a resource they are not even aware of.

Information on Substitute Sites

The distinction that arises over the presence or absence of prespill knowledge is also related to the question of whether a site, such as the Grand Canyon, is "unique" or "irreplaceable," or is one of a large number of similar physical environments. Almost everyone knows about the Grand Canyon, and the problem of close substitutes is unlikely to be important. However, in the case of damage to a "nonunique" site, we need to examine the degree of substitutability across different sites in terms of nonuse preferences. We would expect that any nonuse values are strongly affected by the existence of close substitutes.[27]

Conclusions

Much of the reaction to environmental injury may well arise from considering the possible effects of mankind on the environment in general, rather than on any particular site. This sort of concern is more in the realm of ethical values than in the realm of personal preferences. Just as ethical values concerning income distribution are a basis for government programs to help the poor, but are not a standard part of damage measurement, ethical concerns for the environment are relevant for legislation, but do not fit well in the logical structure of compensatory damages.

HOW CV STUDIES ARE DONE[28]

Survey Methods

The CV method uses surveys in which people are asked how much they are

hypothetically willing to pay (in dollars) to change the condition of some environmental resource.[29] The typical survey begins by giving some background information to respondents on the environmental resource under discussion, then they are told about the change in the environmental condition to be evaluated. They are also told the way money would be collected to finance the environmental change. After this information is presented, respondents are asked for the maximum amount they would be willing to pay to "purchase" the change. Follow-up questions are often used to confirm the answers and to gather demographic information about the respondents.[30]

Thus, the CV method relies on answers to hypothetical questions, such as, How much would you pay if you had the opportunity to purchase the described environmental change? By asking about WTP for the environmental change, the question encompasses both use and nonuse values discussed above. CV practitioners assert that it is important that the environmental change being evaluated be clearly described. For example, individuals might be asked how much they would be willing to pay to preserve a tract of land as wilderness rather than allowing logging of the tract. Or, they might be asked how much they would be willing to pay to prevent the occurrence of some specific number of oil spills of a given size in a given location over the next ten years. The method of collecting the amount to be paid (payment vehicle), if the hypothetical amount were actually to be collected, is also to be described clearly. For example, a surcharge to the federal income tax might be described, or an addition to monthly utility bills. A number of types of payment questions are used in these surveys in attempts to make the question easier to answer or more familiar to respondents. For example, people might be asked if they would pay a stated amount (dichotomous choice) rather than being asked to state how much they would pay. Or, they might be asked to select an amount from a list of amounts on a card (payment card) shown to them. Demographic questions typically include questions about individual characteristics, such as age, education, income level, outdoor activities, membership in environmental organizations, and attitudes regarding the environment.

The answers people give as to how much they are willing to pay to preserve an environmental amenity are interpreted in the context of conventional economic theory. That is, the answer is interpreted as the income change that will leave the individual indifferent between the current situation and the combination of a better environment and a lower income. In economic jargon, the answer is interpreted as a "Hicksian compensated variation."[31] Of course, the existence of a theoretical construct does not imply that the survey questions are in fact measuring this construct.

Evaluating Commodities

In economics, the universally used method of determining the values that individuals hold for different commodities is to observe them purchasing such

commodities and thereby infer that the commodities are worth at least as much as the individuals paid for them. Thus economists have long been skeptical of using the answers to hypothetical questions to assess commodity value, preferring the "revealed preference" approach of relying on observing individual behavior (Samuelson, 1947). This skepticism is owing to the fact that economists (and persons in other fields such as psychology) have only a limited understanding of the motivation of people who are simply asked to answer questions about their preferences. Issues arise over both the care respondents take when they are trying to answer accurately and the reasons they may misstate preferences. It is possible that respondents think they can affect policy, or they may simply want to express their general views. However, when people are observed making purchases for their own use, we assume that they are motivated to spend their own money well. If individual's did not value a commodity at least as much as it cost, they would not be motivated to buy it. In the case of continuously divisible commodities, for example, milk, the last unit purchased has a value to the individual equal to precisely what the individual paid for it. In the case of large discrete purchases, for example, automobiles, we can only infer that the individuals valued the commodity at a level at least as much as the price paid. Of course, individuals sometimes make mistakes in their purchases. So the market is only a fully reliable guide to underlying preferences in the case of repeat purchases, when individuals have acquired experience and understanding in what they are buying, or in the case of those infrequent purchases for which individuals find it worthwhile to learn enough to make an informed decision.[32] It is in the nature of nonuse values that individuals are not observed purchasing such commodities. (Although, in some cases, they are observed making contributions to charity, an issue to which we return below.)

EFFECTS OF CV TECHNIQUES ON RESPONSES

Hypothetical Commodities

A single purchase of an unfamiliar commodity represents a guess as to what the commodity might be worth, rather than an evaluation based on experience. Similarly, hypothetical evaluations of commodities not purchased are also not experience-based reflections of underlying preferences. When experience with the commodity is lacking, considerable thought is required in order to make the response.

People are better able to evaluate hypothetical products if they already have experience with similar existing products. Surveys are widely used to gather information on the feasibility of potential new products. Although these surveys try to determine whether people might be willing to buy different commodities from those currently on the market, survey results are not taken to be one-for-one accurate predictions

of purchasing behavior. Rather, as summarized by Kemp and Maxwell in this volume, these survey results are combined with other information in reaching demand predictions. Experience with the relationship between surveys and actual usage is then used to calibrate the survey responses in predicting demand; however, CV evaluations of preferences for environmental amenities do not make such adjustments.

Moreover, these surveys for potential new products differ from CV surveys in that they are asking people about familiar goods. Nevertheless, such survey research on potential new consumer products has a mixed record in forecasting which products will succeed. The history of American marketing is littered with products for which the suppliers incorrectly thought a sizable market would exist. Indeed, a commonly recited statistic is that about 65% of all new product introductions fail.[33] Even when the commodity is a familiar one, if the nature of the choice is unfamiliar, one can get survey results that are anomalous and suggest the failure of individuals to understand their preferences over the choices available. This outcome is illustrated in a recent article describing a survey on electricity demand in which individuals were asked to respond to questions regarding the value of reliable electricity supply.[34]

In contrast to these two familiar dimensions of commercial surveys, CV surveys ask individuals to evaluate commodities that they do not regularly purchase and may never have thought of as purchasable. Thus both the commodity and the setting are unfamiliar because individuals have no experience with purchasing environmental amenities. They must imagine what such environmental amenities are worth to them in dollars.

Perceptions of Fairness and WTP

Individuals will not assume that ten years without an oil spill can in fact be purchased by the $30 they describe themselves as willing to pay for the commodity; rather, they recognize that the policy under consideration will be financed out of a large number of different pockets. Thus the perceived fairness of who pays the cost among different households and firms in the economy may influence responses. In addition, stated WTP may well vary according to the described method of hypothetically collecting the sums needed to finance the change, whether it is a tax increase, an increase in monthly utility bills, or a rise in the price-per-gallon of gasoline. Indeed, CV practitioners have expressed concern that the choice of a payment vehicle can affect the answers. The plausible sensitivity of hypothetical WTP to the full description of the social situation in which it is implemented suggests that some individuals may be performing a casual social benefit-cost analysis rather than describing preferences over environmental amenities.

The concern about full payment for the cost of environmental changes is separate from another concern in the literature that is based upon survey design considerations, i.e, question framing, or the way in which the question of WTP is asked. As mentioned above, CV surveys have asked for WTP in a variety of different ways. Some CV

surveys use the dichotomous-choice format, in which respondents answer yes-no questions about their willingness to pay a given amount.[35] Other CV surveys ask about individuals' WTP without giving any initial cues about the amount people might be expected to pay; still other CV surveys show people a range of possible answers from which they are to choose an amount. Another dimension of difference among CV surveys is that, in some surveys, respondents are asked a follow-up question to ascertain if their initial response really indicated the maximum amount they would be willing to pay. The literature suggests that different ways of asking these questions generate different answers.

Self-Expression versus WTP

Because of the hypothetical nature of the questionnaire, it is also plausible that some individuals are expressing a concern about a larger set of environmental changes or a general environmental issue rather than describing a genuine willingness to pay that would be followed up by an obligation to pay.[36] Although the question requires that interviewees express their degree of concern in dollars, expressing a concern in dollars need not be the same thing as recognizing a genuine willingness to pay the same number of dollars. Thus, respondents' stated WTP may reveal more about their attitudes or intentions than about their actual behavior.[37] As economists sometimes note, talk, in response to surveys, is cheap; actual behavior costs money.

Questionnaire Design

A natural concern about the reliability of the answers to CV surveys arises when the answers vary a great deal when there are apparently small differences in the design and implementation of CV questionnaires. It appears that differences in question wording, differences in question sequencing, and differences in individual interviewers can all have significant impacts on the range of answers.[38] An illuminating example comes from a CV study by Samples and Hollyer (1990) concerning the value of preserving seals and whales. Some respondents were asked the seal question first and then asked the whale question. Other respondents were asked the questions in reverse order. Seal value tended to be lower when the seal question was asked after the whale question, whereas whale value was not affected by the sequence of questions. The authors offer an explanation, which was based on a debriefing of one of their interviewers:

> Apparently, when respondents valued seals first, they used their behavior in this market situation to guide their responses to whale valuation questions. Since whales are generally more popular than seals, respondents were reluctant to behave more benevolently toward seals compared with humpback whales. Consequently, whale values were inflated in the S-W questionnaire version to

> maintain a relatively higher value for the humpbacks. This behavioral anchoring effect did not exist in the W-S version, where whales were valued first. (p.189)

The literature has focused on which method is the "right" way to ask the question in order to maximize precision. But when the results are sensitive to the question format, are the answers related to underlying economic preferences, as the CV method assumes? These specific concerns about the right question format lead to a broader concern: Can the underlying preferences be measured without bias by any of the formats? No other method exists for measuring nonuse values. Therefore, there is no standard against which CV answers can be compared to detect bias.[39] So, we stress the concern of survey bias, not the concern of increasing survey precision.

In the absence of external checks on the correctness of answers to CV surveys, one needs to search for internal consistency checks as a minimum test for the validity of the survey answers. However, one also needs to recognize that systematic bias has its own logic and consistency. Thus, hypothetical WTP responses might satisfy some consistency checks without the answers equaling the values being sought. Therefore, we view internal consistency checks as a minimum standard, but they cannot validate the overall process in the absence of real-world benchmarks.[40]

REFERENDA

We do observe individuals making decisions that affect their economic well-being in situations other than market situations. For instance, individuals sometimes vote on binding referenda that affect both the taxes they pay and government expenditures. Thus, a referendum can be considered an analogy to the attempts to measure nonuse values through questionnaires.[41] But, does voting successfully elicit economic preferences, even when the outcome is binding (as opposed to a nonbinding referendum "to send a message" to politicians)? Considerable skepticism exists about the extent to which voting on a referendum represents informed decision making.[42]

It is also interesting to consider polls about voting intentions before such referenda. Sometimes, polls are accurate predictors of voting outcomes. Sometimes, polls are not accurate predictors, even when they are taken close to election day. What is particularly interesting is to consider repeated polls about the same referendum. Magleby (1984) has analyzed statewide polls in California and Massachusetts for which there were at least three separate surveys. In some cases, the polls show roughly the same margin over time. Magleby calls these "standing opinions" and believes that this stability comes from the voters' deep attachment to their opinions on some controversial issues, such as the death penalty and the equal rights amendment. In some cases, the polls show significant changes in the margin of preferences but no change in the side that is ahead. Magleby calls these

"uncertain opinions." Examples of such votes involve requiring handgun registration and permitting homosexuals to teach. In some cases, there are significant changes in voting intentions as the campaign proceeds, with victory in the actual election going to the side that had at one time been far behind. Magleby calls these "opinion reversals." Examples of such votes are a state lottery, a tax reduction measure, and flat rate electricity. In the latter case, a February poll showed 71% in favor, 17% opposed, and 12% undecided. The actual vote was 23% in favor and 69% opposed. Seven percent of the voters skipped this question. In his analysis of 36 proposition votes in California, Magleby found 28% were standing opinions, 19% were uncertain opinions, and 53% were opinion reversals. That is, in a majority of cases, early opinion polls were not good predictors of election outcomes. Moreover, they were not even good predictors of later opinion polls, after the campaign had run for some time.

A very recent example of the inaccuracy of opinion polls with respect to environmental issues is the vote on the "Big Green" Proposition 128 that appeared on the California ballot in 1990. The Big Green initiative would have taxed oil companies to pay for potential oil spills, saved old-growth redwoods, reduced gases claimed to cause global warming, and banned suspected cancer-causing chemicals in food. A June opinion poll reported in the *Los Angeles Times* found that 55% of respondents who had an opinion (overall, 84% of respondents had an opinion) stated they were in favor of Proposition 128. The week before the election, the *L.A. Times* opinion poll found equal voting intentions for and against Proposition 128. Yet, at the actual election, Proposition 128 lost badly, getting only 36% of the vote. Supporters of Proposition 128 claimed that an advertising campaign by its opponents caused voters to change their minds. Whatever the reason for the large difference between the opinion polls and the actual votes, the experience of Proposition 128 demonstrates that individuals do react to the level of information they have and the context of the situation (e.g., many [28] actual propositions on the California ballot) in which they make decisions. Thus, the referendum model claim for the use of CV has no foundation in individual economic preferences.

It seems to us that responses to CV questionnaires for a single environmental issue are likely to be based on little information because of the reasons we have outlined above. This suggests that the outcomes of CV surveys are not likely to be accurate predictors of informed opinion when people are given more time to reflect on a particular issue and are given more information about it. Such surveys are therefore unlikely to be a good basis for either informed policy making or accurate damage assessment. In the functioning of a democracy, the quality of decision making by voters may not be as important as the placing of some powers directly with the voters. In contrast, WTP answers must be accurate expressions of economic preferences if the CV method is to be used in administrative decisions and judicial proceedings.

CRITICISMS AND INCONSISTENCIES OF CV

Above, we raised a number of questions about the CV survey approach, which attempts to measure total values, and so nonuse values, in responses to hypothetical questions. In this section, we focus on the issues raised by the results of such surveys. Because no alternative direct measure of preferences exists for nonuse values, we do not have a direct and conclusive way of comparing the answers generated by CV surveys with the "correct measure." Thus we are reduced to considering indirect tests to verify that CV is properly measuring preferences. A number of different partial tests can help determine if the answers provided by CV surveys are satisfactory. One place to start is with the interviews and debriefings of people actually filling out CV surveys.

Verbal-Protocol Study

A verbal-protocol study asks individuals to "think aloud" while answering questions. In responding to a CV survey designed to elicit their willingness to pay to protect migratory waterfowl from drowning in waste-oil holding ponds, individuals were asked to think aloud as they responded to the questionnaire and to report everything that went through their minds.[43] Everything they said was recorded on audio tape, and these tapes were transcribed and coded for the type of considerations being mentioned. Respondents verbalized many diverse considerations. Perhaps the most common strategy involved first acknowledging that something should be done and then trying to figure out an appropriate payment amount. About one-fourth of the sample mentioned the idea that, if each household did its part, each household would not have to give all that much. About one-sixth of the sample attempted to estimate how much they personally would be affected by a change in gasoline prices that would finance the cost of protecting the waterfowl. For example, some respondents converted an increase of a nickel a gallon for gasoline into an annual expenditure. Neither those who felt all people should do their part nor those who considered the effect of the gasoline price increase were attempting to value the underlying good. About one-sixth of the sample made comparisons with donations to charities. (We argue below that experience with actual donations to environmental charities is not closely related to evaluating the services provided by the charities.) About one-fifth of the sample said they just made up a number or guessed an answer. This result may reflect the unfamiliarity of the task they faced. Many respondents seemed to wish to signal concern for a larger environmental issue. These findings strongly suggest that people are not easily in touch with underlying preferences.

Implausible Responses

Another element that suggests the surveys do not measure preferences is the

presence of some very large WTP responses that seem implausible. When individuals say that they are willing to pay several thousand dollars, despite not being very wealthy, their hypothetical answers suggest not only that the particular answers don't make sense, but also that the questionnaire is not an adequate device for evaluating underlying preferences. The CV literature has long recognized the problem created by these extreme values, or "outliers." Standard practice has been to eliminate (or down weight) some of these outliers in order to generate more plausible numbers. But, instead of thinking of the problem as one of statistical inference, we could view the presence of a sizable number of implausible answers as evidence of the likelihood of bias or incorrectness in the answers generally. Similarly, the presence of large numbers of "protest" zeros suggests a difficulty in answering the question being asked[44]—as does the presence of a significant fraction of "don't know" answers in some surveys.

Evidence frequently cited in support of the plausibility of CV answers comes from the consistency of CV answers with empirical experiments for the same goods. Proponents of CV often cite the "strawberries" experiment[45] and the "goose hunting permit" experiments.[46] In both of these cases, people were valuing familiar commodities they use, not unfamiliar environmental sites they may never use. Also, people were engaged in a transaction familiar to them, rather than an unfamiliar one, such as valuing the environment. Thus, the fact that CV works somewhat for standard goods does not mean that CV works for public goods.

Follow-Up Questions in CV Surveys

Further evidence suggests that individuals are not really describing preferences over natural resources in CV surveys. Follow-up questions in surveys ask people whether their answers reflect only their WTP for the particular resource being valued or are, instead, intended as a general contribution to environmental causes. If the respondent is considering the general environment, CV proponents assume that the respondent can accurately separate the value of the specific resource from a general contribution. (In fact, a sizable fraction of individuals respond that their answers indeed include contributions to general environmental causes).[47] Our interpretation of answers to follow-up questions is that people are confused about the nature of the questionnaire and do not distinguish between a particular resource and the general environment.

Direct Test of Charitable Giving

Two studies relate WTP responses to actual charitable donations. In a study in Norway (Seip and Strand, 1991), a sample of individuals were asked how much

they were willing to pay for membership in the most important environmentalist organization in Norway. Those individuals stating a WTP equal to or above the actual membership fee were sent brochures and an invitation to join the organization. Only 6 out of 64 respondents who initially expressed a WTP in excess of the membership fee in fact joined the organization.[48]

In a study in Montana (Duffield and Patterson, 1992), two samples of fishing license holders were selected and sent a mail survey.[49] One group was asked for their hypothetical willingness to pay to support the Montana Nature Conservancy, which wanted to buy water rights in order to have greater water flows in two rivers where they might fish. The second group was given the same information and solicited to actually donate to the Montana Nature Conservancy for the same purpose. Among Montana residents, 33.2% of respondents to the hypothetical survey said they were willing to contribute. Given the response rate, this proportion was 6.3% of deliverable surveys. Yet only 1.1% of deliverable requests for cash resulted in cash payments. Similarly, 56.3% of nonresident respondents stated a positive WTP. This proportion represented 15.4% of deliverables; yet, only 5.7% of deliverable requests for cash resulted in payments.[50] Thus, in both of these experiments, there was a clear mismatch between the hypothetical CV answers and the actual WTP. More generally, an enormous mismatch exists between the values derived from answers to CV surveys for protecting the environment and actual donations to environmental causes. As summarized in the paper by Mead in this volume, it is common to find the hypothetical WTP to support an environmental issue in excess of one, or even ten, billion dollars. Yet, in 1990, total giving to environmental and wildlife causes was estimated to be $2.29 billion.[51] Among the three-quarters of the U. S. population who made some charitable contribution in 1990, only 17.8% donated to environmental causes.[52] We also know that only a small fraction of the population contributes when given the opportunity to donate to help protect wildlife through tax check-offs on state income taxes.[53] Moreover, these contributions are far more likely to be made by people receiving rebates than by people paying additional taxes, suggesting that mood, rather than preference, plays a large role in such behavior.[54]

It is natural to consider reasons why charitable donations might be significantly less than WTP for environmental goods. The obvious explanation for this difference is "free-riding," a term that describes a situation in which individuals who care about the environment do not contribute to environmental causes, such as the Nature Conservancy, which buys land to prevent development. They expect other people to contribute enough, so that, at the margin, an additional dollar given to such a cause is worth less than a dollar to the individual potential donor. Although we have no doubt that free-riding is an important element in individual behavior, direct empirical evidence on free-riding shows that in experimental settings people behave neither like the altruist, who gives correct preference evaluations, nor like *homo economicus*, who exhibits purely self-interested behavior.[55] Rather, studies find a variety of behavior.

In analyzing how valid CV is, we must come to grips with free-riding. If

free-riding is not a common trait, then it is impossible to reconcile charitable donations with taking literally answers derived from CV surveys. If free-riding is the reason for the large difference, then one would expect that similar strategic behavior would contaminate CV responses. It is a common assertion of CV practitioners that little evidence of significant strategic behavior has been observed in responses to CV surveys.[56] It seems plausible to us that, if CV surveys are in fact measuring preferences accurately, without the presence of strategic behavior, then the same behavior should carry over to actual charitable contributions. Implicitly, CV practitioners must argue that individuals do not behave strategically when their money is not on the line, but do behave strategically when their money is on the line. If this sharp difference between actual preferences and answers to hypothetical questions is brought out by the hypothetical nature of the CV survey, it is equally plausible that other sharp differences between actual and hypothetical answers as well must be recognized and taken into account.

Willingness To Pay and Willingness To Accept

Consider the issue of visibility at the Grand Canyon, recognizing how visibility varies throughout the year. Consider a costly project that can decrease pollution from power plants and thus improve visibility on some of those days. Next, consider a CV survey that asks respondents how much they are willing to pay to fund this project to improve visibility. Instead of this survey, consider an alternative survey in which the respondents are told that the costly project has actually been approved (rather than just being proposed). Then tell the respondents that the government is considering saving money by canceling the project. In this alternative survey, the respondents are asked a *willingness-to-accept* (WTA) question: How much money would the respondents have to receive to be in favor of canceling the project (thereby accepting worse visibility)?

The two questions involve the same change in visibility. Thus one might reason that the two questions should receive the same answer, but, in fact, CV studies frequently find that WTA greatly exceeds WTP. A possible explanation for this difference is the "income effect." Respondents to the WTP question face a situation in which they start with poor visibility and must pay to improve visibility (by funding the project), whereas the WTA question respondents begin with the higher visibility level (at no cost to them). In effect, the less desirable initial situation faced by the WTP respondents is equivalent to having a lower starting income. Because expenditures for most goods, including environmental resources, decrease when income is lower, one would expect WTP responses to be lower than WTA responses. Nonetheless, the size of this income effect should be quite small for two reasons. First, the money value of the difference in initial well-being between the WTP and WTA scenarios is small. Second, CV surveys have shown that stated WTP does not increase in proportion to income.[57]

CV proponents offer two responses to the finding of large differences between WTP and WTA. One response is to "be conservative" and take the lower (WTP) amount, a response that ignores the fact that the persistent difference between theory and observation casts doubt on the whole enterprise. The second approach is to argue that such large differences are, in fact, consistent with economic theory (Hanemann, 1991). Of course, CV practitioners need to argue not just that the results are conceivably correct, but that they are indeed plausible for the kind of resources in question.

We can conceive of hypothetical questions for which we would expect enormously different WTP and WTA amounts. When asked how much you are willing to pay to avoid instant death, you would probably answer all your available wealth. On the other hand, the amount you would need to be paid in order to accept instant death is probably infinite. Thus, it is certainly possible to have preferences that would lead to large differences between WTA and WTP. But, does this apply to the situation of particular environmental resources?[58] After all, is the risk of an oil spill in a remote area qualitatively similar to having one's life in danger? In the latter case, your WTP is very large, relative to your income. But CV measurements of WTP for environmental protection are quite small relative to income. Thus it seems more likely that the difference between WTA and WTP is yet another sign of the failure of CV to measure correctly any genuine underlying preferences for nonuse values.

Further, we can consider the argument by Opaluch and Grigalunas (1991) who start with the premise that, with environmental protection at issue, we get primarily ethical answers rather than true preferences. From this point of view, WTP indicates a willingness to give up money for the pursuit of one's ethical values. On the other hand, WTA is ethically analogous to accepting a bribe and so is enormously larger or infinite. This reasoning is consistent with the common phenomenon of some individuals refusing to answer questions with such ethical overtones. Thus, CV answers are seen as expressions of ethical values, not descriptions of preferences.

Sequence of Questions

A long-standing anomaly in CV evaluations has been that answers depend on the sequence of other evaluations that respondents are asked to make. Thus, Schulze et al. (1983) found a sample of Chicago residents willing to pay $90 to improve visibility at the Grand Canyon. Yet, in a study one year later, Tolley et al. (1983) found that a sample of Chicago residents was only willing to pay $16 for that visibility increase at the Grand Canyon. Why the difference? Because the second sample of residents had been asked first to state WTP for improved visibility in Chicago and for increased visibility in the rest of the East.[59] Could $90 and $16 both be the correct answer? Two standard defenses of the inconsistency of

these answers have been advanced by CV proponents. The first defense is based on income effects, and the other is based on substitution effects.

The defense based on income effects is that, having hypothetically spent money for visibility improvements elsewhere, you have less income to spend for additional visibility. For us to accept this explanation, the required income effect would have to be enormous. Yet, when we look at how people with different incomes answer WTP questions, we see that WTP does not increase drastically with income: in other words, income effects are small.[60] Thus, the income effect defense of CV proponents is contradicted by the results in CV surveys.

CV proponents then defend the anomalous results ($90 versus $16 for Grand Canyon visibility) by citing substitution effects. They argue that improvements in visibility in the eastern half of the country are substitutes for improved visibility at the Grand Canyon. They then argue that this substitution is sufficiently powerful to reduce the WTP from $90 to $16. To test this argument, we could first tell respondents that visibility will be improved in the eastern U.S.; then we would ask them how much they would be willing to pay for improvements in the Grand Canyon. If the substitution effect argument is correct, then we should get around $16 as an answer, the same number as in the Tolley et al. (1983) study. However, our intuition is that we would get around $90 as an answer. Such a finding would imply that the substitution effects argument is incorrect.

Wilderness Study

We decided to test the consistency of CV responses directly by asking different samples of individuals different questions, permitting us to test directly the substitution effects in CV responses.[61] The purpose of the wilderness CV experiment was to examine how CV answers for protecting a wilderness area vary with the number of threatened wilderness areas.

For this purpose, CV surveys were administered to people living in four western states that contain 57 different federally designated wilderness areas. People were first informed that a certain number of these areas were going to be opened up for logging to generate revenue for the government. Then, people were asked their WTP to protect an additional area from logging.[62] Different surveys asked about different areas. In addition, the number of threatened areas was different across the surveys. We wanted to see how people valued an eighth area after being told that seven areas were being threatened. We were then able to compare how they valued a ninth area after being told eight were being threatened. If, in fact, there were large substitution effects, we should have found large differences between the answers. Instead, we found that the answers were not different from each other, according to a statistical test.

The wilderness CV survey then asked people about WTP to protect three areas at once. The answer to this question was different from what we got when

we added up the answers for valuing the three areas individually (according to a statistical test). In fact, the answer for preserving three areas was not different from the answer to protect one area (according to a statistical test). Given the absence of a substitution effect (for which we had already tested) and the lack of large income effects in the answers, we conclude from the wilderness CV experiment that the answers to these questions are not consistent with underlying preferences. Two different measures of the same preferences yield different (and, so, inconsistent) answers. Without consistency, the consumer choice axiom of transitivity cannot be satisfied. So, we conclude that responses to CV questions are not consistent with the basic economic theory of choice.

In another type of CV survey, individuals are asked two yes-no questions. First they are asked whether they are willing to pay a given amount (say $20) to protect an environmental amenity. If they say yes, they are then asked whether they are willing to pay a larger amount (twice as much, or $40) for the same improvement. On the other hand, if they say no to the first question, they are then asked whether they are willing to pay a smaller amount (half as much, or $10) for the same improvement. The survey can be designed so that some people are asked about their willingness to pay a given amount as a first question while others are asked about the same amount as a second question. If the answer to a $20 question is not affected by whether it is first or second in the sequence, the fraction of "yes" answers should be the same in both settings, subject to sampling variation. In a paper in this volume, McFadden and Leonard reject the hypothesis that the answers are consistent with the theory of economic choice.

Waterfowl and Oil-Spill Studies

In 1986, Kahneman reported that the stated WTP of Toronto residents to prevent a decline in fish populations in all Ontario lakes was not appropriately greater than the stated WTP to prevent a decline in the lakes in a small area of the province.[63] To document the robustness of such a failure of reported WTP to vary appropriately with the magnitude of the environmental improvement, a similar study was done in Atlanta.[64] This study examined stated WTP to save migratory waterfowl, some of whom die each year in uncovered ponds containing wastewater, oil, and other by-products from oil and gas drilling operations. One sample was asked about saving 2,000 birds; another sample, 20,000 birds; and a third sample, 200,000 birds. The distributions of responses were not statistically different across the three samples.

Similarly, two different samples were asked for their WTP to finance oil-spill response centers. One sample was asked for WTP to have local response centers that would be located at ports to handle small spills. The second sample was asked for WTP to finance these centers and also regional centers that would handle larger spills. This time there was a statistically significant difference between the responses in the two samples, but the small-spills responses were higher than the all-spills

responses.[65] Again, the conclusion is that responses to CV questions are not consistent with the basic economic theory of choice.

Top-Down Disaggregation Study

As mentioned above in the discussion of the wilderness study, the stated WTP for three areas together was much less than the sum of WTP's for three areas valued individually. Conversely, we might focus on the implications of disaggregating the stated WTP for all three together into three separate parts corresponding to the three areas.

Kahneman and Knetsch (1992a) conducted an experiment to explore this type of disaggregation.[66] In their CV experiment, they asked people how much they would contribute to a special fund for improved environmental services, with the proviso that money collected would lead to significant improvements. Note that this form of questionnaire is somewhat different from the ideal CV survey of a public good, in which individuals are offered a specifically defined good and asked for their willingness to pay for that good. The standard form of the question offers the respondent the opportunity to purchase an improvement in the environment (Grand Canyon visibility). The Kahneman and Knetsch form of the question, however, is like an offer to contribute to an environmental charity. Individuals are not purchasing a specific environmental improvement; they are contributing to a cause that works toward general environmental improvement.

We would expect that if people can contribute to cause A or to cause B, or, as a third option, to a fund C that covers both causes A and B, their contribution to fund C should be nearly the same as their contributions to causes A and B combined.[67] But Kahneman and Knetsch found that when individuals were asked how much they would give to a single cause, they stated much larger amounts than they would have allocated to the same cause had they first been asked how much they would contribute to an array of causes and then asked to divide that contribution among the different causes. Kahneman and Knetsch interpret this result as evidence that the answers to CV surveys are not an accurate reflection of underlying preferences. They conclude, instead, that responses to CV questions are not a reflection of preferences over resources but a purchase of moral satisfaction obtained by supporting good causes.[68]

In an attempt to evaluate just how wide a range of answers one can get from different degrees of embedding, Kemp and Maxwell completed systematic surveys involving different degrees of embedding[69] and found ratios of several hundred to one in answers to essentially the same question posed in different ways. This magnitude of difference underscores the implausibility of the claim that answers in CV surveys reflect underlying preferences. Whereas the wilderness study attempted to measure exactly the same preferences in two different ways, and showed that it is impossible

that the CV answers reflected preferences, the disaggregation studies showed that it is implausible that the respondents' CV answers reflected preferences.

A THEORY OF POSITIVE RESPONSES TO CV SURVEYS

Picture yourself as a typical member of the American public, not as a professional economist or policy analyst. Picture yourself being asked to decide on the amount it is worth to you to have some environmental amenity protected. Assume that it is an amenity you have never seen or used nor expect to see or use. Assume that you are trying hard to cooperate with the survey questionnaire and trying to provide a meaningful answer for whatever purpose the survey designer has in mind. To what would you refer in your experience or intuition in order to generate an answer? For a large fraction of people asked CV questions the answer is zero; they do not choose to express a positive WTP to protect an amenity under these circumstances. A large proportion of zeros is a common finding in CV surveys.[70] In the surveys reported above, about preventing logging in federal wilderness areas, on average, about 40% of the respondents answered zero. In the usual follow-up question to find out the reasons for zeros, many describe themselves as poor, whereas others describe themselves as in favor of logging. Some people express a lack of concern for the wilderness. With a request to provide a rationalization for the answer of zero, some individuals are pressed to think up a suitable reason. Thus it is not clear how seriously one should take these answers.

But what about the people who do give positive answers? Given the large fraction of zeros, we might expect a large number of very small answers: $.20, $.30, $.50, $1. In fact, very few such small answers are typically found in CV surveys. The distribution of answers for a typical CV survey has a large number of zeros and a large number of answers in sizable round numbers (in technical terms, the distribution is bimodal). People who decide to report a positive answer must generate some decision rule for finding an answer. This need shows up in the verbal-protocol study mentioned above. For some people, the process might be to consider how an amount per person converts into an aggregate that seems suitable to finance the change. For some, the process might be to select a small change in the payment vehicle, such as an extra penny per gallon for gasoline, and then to convert (perhaps inaccurately) this seemingly small amount into an answer by multiplying by the number of gallons of gasoline typically purchased.

Other respondents explicitly recognize the parallel between typical CV questions and requests for charitable donations. These respondents' answers might reflect how they would have dealt with the following situation: Assume that, of the various environmental charities to which you might contribute, you have decided to contribute to this one. How much would you then contribute? The explicit recognition of the

parallel between charitable contributions and WTP in some of the verbal protocols support our consideration that respondents' actual charitable contributions play a significant role in our understanding of answers to CV surveys.

Some charities attempt to collect small change in many locations. Other charities solicit individuals directly. For the latter charities, we see the same pattern in gifts—many zeros and some sizable checks (a bimodal distribution).

The "Warm Glow" Effect

As we will argue below, the evidence on charitable giving is consistent with the "warm glow" that comes from the act of giving per se—in addition to what is accomplished by the resources that are transferred. This "warm glow" theory of charitable donations is further support for the idea that CV does not measure underlying preferences. If people give charitable contributions in large part for the pleasure of giving,[71] then it is also plausible that they will give large answers to CV questions asking them to evaluate environmental improvements in dollar terms. This reasoning implies that the answers do not reflect an evaluation of the resource being protected. The warm glow from actual giving is obtained by making a small number of relatively large donations. For hypothetical contributions, one can achieve a "warm glow" from an unlimited number of large answers.

Given the "warm glow" effect, it is interesting to review the findings reported above on the survey of WTP to preserve wilderness areas. We find that the stated WTP is roughly the same to protect only one wilderness area or to protect three. This pattern is consistent with a warm glow interpretation of a situation in which the degree of warm glow does not vary much with the relatively small differences in scenarios, both in terms of a small number of areas being protected and a small difference in the number of other areas threatened. Given our view that the CV responses resemble answers about the amounts that one might give as charitable contributions, we tested to see if preserving all 57 areas would lead to a statistically significant different answer. It does indeed.[72] When individuals were offered a substantially larger cause, the stated WTP of those giving positive answers was significantly larger, and the proportion giving positive answers was a bit higher. Our conclusion is that people do indeed care about preserving wilderness areas, but we infer that standard CV questionnaires do not generate a description of preferences but, rather, elicit responses that generally express concern about preserving wilderness. One can interpret this result as saying that the warm glow effect makes up the overwhelming fraction of the answer; that is, the relationship between CV answers and true preferences about the public good of wilderness preservation is not remotely close to one-to-one and is unknown. Thus the answers generated by CV surveys are not reliable guides to underlying preferences over environmental resources.

Charitable Giving

To place CV evaluations in context, it seems useful to review some of what is known about charitable giving. Let's consider four observations about charitable donations. First, aggregate charitable contributions to all environmental causes are small relative to CV evaluations. Second, surveys of whether individuals have given to charitable causes are often inaccurate. Third, evidence suggests that people do not have a good awareness of what is being accomplished by their donations. And fourth, significant evidence has been found that a great deal of the motivation for charitable giving comes from the act of giving rather than from the "commodities purchased." We have already cited some of the statistics on the small aggregate contributions to environmental charities and the small fraction of the population that contributes through tax check-offs to wildlife funds. As we argued above, when you combine this level of actual charitable donations with the high level of hypothetical contributions for environmental resources, it suggests that the hypothetical responses are not accurate reflections of true WTP.

In the survey of WTP for protection of wilderness areas described above, individuals were asked in the follow-up questions if they contribute to organizations that actively work to preserve and protect nature or wildlife. Consistently, 35 to 40% of the surveyed individuals answered yes. Although this sample was not a random sample of the entire U. S., the response proportion is unlikely to be consistent with the estimate that less than 15% of the population contributes to such organizations.[73]

In one of the early attempts to evaluate the accuracy of answers to surveys, Parry and Crossley (1950) examined the accuracy of reported donations to the Denver Community Chest. They found considerable exaggeration of the extent to which people had donated.[74] This finding not only raises questions about surveys, but also underlines the social pressures associated with both the act of giving and with the act of answering questions about giving. If the former did not exist, the latter would not be so important.

In a survey of Indiana donations to United Way, Keating, Pitts, and Appel (1981) wrote:

> On average, a respondent could name only one of the 34 funded agencies. In addition, many respondents named agencies which were not funded by United Way . . . Since the ability to merely name funded agencies is low, we expect that any in-depth knowledge of agency service or effectiveness is virtually nonexistent. (p. 821)

The relevance of this finding is that much of actual giving is not based on detailed evaluations of commodities. It is therefore not surprising that people might respond to CV questions without having an evaluation of the environmental amenity about which they are asked.

People often complain about social pressure to give to charities.[75] Studies of the role of social pressure in generating charitable contributions generally find

statistically significant effects of the nature of the solicitation.[76] This combination of results underscores again the role of the act of giving, per se, in charitable behavior. Although some parallel with advertising for ordinary commodities might be found, the parallel is limited.

The economics literature on charitable contributions has concentrated on the relationship between total contributions of an individual and income and tax deductibility. An important puzzle in behavior toward charitable contributions is seen in the behavior of individuals faced with a choice of a number of different environmental charities.[77] If your preferences for charitable giving were determined solely by the use that would be made of the money, and if your contributions were small relative to the total budget of the organization, then the logic of consumer choice would imply that you would concentrate all your donations on a single cause. (Your thousandth dollar would do as much "good" as your first dollar, so you would not want to switch to another charity.) The fact that people spread their donations over a considerable number of similar charities strongly suggests that satisfaction is associated with the act of donating. However, because you probably get satisfaction from donations that are large compared to the size of the average donation, you likely will contribute to a considerable number, but not all, of the charities at hand. You will tend to make small donations to those expecting change and larger donations to those expecting checks. Moreover, you are likely to choose representative charities from a variety of different types of charities.

These conclusions help to explain our interpretation of responses to CV questions. Donors might choose one or another charity and then give a sum to the chosen charity, giving nothing to the second. We think it plausible that some of the responses in CV surveys reflect two questions: "Is this cause the type of cause to which you might donate?" and, "If the answer is yes, how much would you give if this cause were the charity you selected from the charities in the acceptable group?" Some of the positive responses are likely to be responses to the question, "How much would you give if you selected this cause as your environmental charity?"

CV IS INAPPROPRIATE FOR BENEFIT-COST ANALYSIS AND ECONOMIC DAMAGES

The central argument we have made is that the respondents' answers to CV questions do not arise from an examination and expression of underlying preferences. Because answers to CV surveys do not measure preferences, they are not a suitable source of information on values in benefit-cost analysis. In addition, CV surveys cannot be used to estimate the loss that people have suffered and, therefore, are not suitable to measure compensatory damages.

Many aspects of the natural environment and man-made impacts on it are probably not well understood by the general public. For example, the general public is probably not well informed about the wide range of natural variation in some species populations, nor about the temporary nature of much of the damage from oil spills. It makes no more sense to rely directly on ill-informed members of the public to evaluate the dollar value of such environmental damage than it would be to rely on an ill-informed public to choose between alternative designs for airplanes or nuclear power plants.[78]

A natural tension arises in representative democracies between lodging complicated decisions in the hands of the general public (through referenda) or having elected representatives rely on the combination of expert advice and the political process. As an issue increases in complexity, the amount of knowledge required for an informed opinion increases—and the ability of the general public to make good decisions decreases. Our findings above imply that CV-generated answers are more like public-opinion polls than like scientific or economic evaluations of underlying realities. For nonuse values, CV surveys focus specifically on questions with which people have little or no experience because they generally do not use the resource. We conclude that CV surveys are not a reliable basis for sound decision making.[79]

What if fines were based on the CV answers reported above? Stated WTP did not vary according to the number of birds saved by covering waste-oil ponds. If fines for killed birds were used as an incentive for protecting migratory waterfowl, the fines would not vary with the number of killed birds over a considerable range. This system is surely not an efficient way to create incentives for wildlife protection. Likewise, we found that the WTP responses for protecting one or two wilderness areas were similar. A developer proposing two projects could then argue that blocking one project would be sufficient protection of the wilderness, with the public placing no value on blocking a second project. We think that an efficient method for protecting wilderness would involve shadow prices on wilderness areas that varied with the land preserved, rather than shadow prices that are insensitive to the amount of the resource in question. Moreover, the presence of warm glow makes CV-based shadow prices dependent on the number of proposed developments that don't happen. Efficiency calls for shadow prices that only depend on what does happen.

Because there is no clear link between the results of CV surveys and underlying preferences, no clear basis exists for selecting among different ways of doing CV surveys. Yet the results of CV surveys are extremely sensitive to the choice of how a survey is done and interpreted.[80] Thus the desire to have a scientific basis for nonuse values by looking to preferences becomes, instead, a vehicle for manipulation of answers by the choice of methods.

CV does not provide a reliable method to calculate natural resource damages. The inevitable outcome is great uncertainty about the level of damages that may be assessed because a clearly defined correct method of doing CV evaluations is only a figment of CV proponents' imaginations. This uncertainty in legal outcomes

creates uncertainty for firms engaged in activities that are possibly subject to a lawsuit involving nonuse value. This uncertainty is likely to be a source of economic inefficiency.[81] In traditional damage measurement, uncertain, subjective components of loss are normally excluded from damages.

If CV is deemed reliable in measuring nonuse environmental damages, what is to stop its application to many other circumstances? Consider the highly publicized death of a child resulting from a highway accident caused by a combination of poor driving and poor road design. Suppose people are asked how much they would contribute to improving the road to avoid such a death for the next ten years. It is likely that the answers would be large. Would it make sense to assess damages for poor driving that reflected such a large sum? Because a person would be bankrupted by such a CV damage estimate, people would want to insure themselves for billions of dollars before they would choose to drive. It is ludicrous to suggest that CV would be an appropriate method for awarding damages. Similarly, using CV to do benefit-cost analysis in this situation would call for spending vast sums to increase highway safety and possibly would call for dramatically reduced speed limits as a way to save lives. Such an allocation of resources would be inappropriate.[82]

Concern exists that the use of only economic calculations based on actual market evaluations for benefit-cost analysis and compensatory damage measurements might leave the environment in a worse state and at more risk than it would be with a good political process. If nonuse values cannot be measured, what other approaches can take account of the environmental concerns? The political process has and can legislate to protect the environment. Legislation for government spending to protect the environment and for penalties and rules associated with risks to the environment need not rely only on conventional economic values; it can also reflect public concerns for the environment. An attempt by some economists to use implausible and unreliable findings is inappropriate as a response to environmental concerns.

NOTES

[1] This paper reports on research funded by Exxon Company, U.S.A. The opinions expressed are those of the authors and not necessarily those of Exxon.

[2] The revealed preference technique of economic analysis was pioneered by P.A. Samuelson in the 1930s.

[3] For example, in a "conservative" damage assessment of the *Nestucca* oil spill, Rowe et al. (1991) found that residents in Washington state stated they were willing to pay from $65 to $175 per household to prevent such a spill. Rubin, Helfand, and Loomis (1991) found an aggregate WTP of $1.48 billion per year to preserve the northern spotted owl. For more examples, see the paper by Mead in this volume.

[4] For presentations and discussions of nonuse value and the CV method see Cummings, Brookshire, and Schulze (1986) and Mitchell and Carson (1989).

[5] Option value was introduced in this context by Weisbrod (1964). As we indicate below, existence value, introduced by Krutilla, is the primary focus of CV evaluations.

[6] Footnote in original: "Uniqueness need not be absolute for the following arguments to hold. It may be, like Dupuit's bridge, a good with no adequate substitutes in the 'natural' market area of its principal clientele, while possibly being replicated in other market areas to which the clientele in question has no access for all practical purposes."

[7] This is not to say that people don't make mistakes. Reliability of market behavior as a guide to underlying preferences is clearly greater when people are familiar with the good they are purchasing than when they are not.

[8] The methodology is similar to examining the cost of purchasing a good (its price) in contrasting the positions of having a good and not having it. Thus, measurement of lost use values is based on inferences that arise from individual actions or changes in these actions.

[9] For the examination of foregone use, see the paper in this volume by Hausman, Leonard, and McFadden.

[10] See Weisbrod (1964).

[11] Option value as used here corresponds to options that are traded on stock exchanges. A stock option gives an individual the opportunity, but not the obligation, to purchase 100 shares of a given stock at a given price over a specified time period in the future.

[12] See Schmalensee (1972).

[13] The small size of option values arises because the use values for a particular site are typically small in relation to an individual's overall budget and because of the existence of (close) substitutes for the given sites under evaluation.

[14] See Arrow and Fisher (1974).

[15] See, for instance, Freeman (1984).

[16] Economics often combines the preferences of parents and children as if a family behaved in the same way as a longer-lived person who is not married. Thus, bequest value recognizes that parents may wish a natural resource to be available for their children.

[17] Below, we will consider existence value, a value unrelated to human use. Bequest value includes the bequest of existence values as well as use values.

[18] For a discussion of the connection between altruism and the proper design of benefit-cost analysis, see the paper by Milgrom in this volume.

[19] Externality values can be either positive, for example, sharing the experience, or negative, for example, congestion.

[20] For a discussion of earlier attempts to separate out the different values discussed above, see Cummings and Harrison (1992). The authors are quite skeptical that earlier studies succeeded in such attempted separations.

[21] We recognize that a large number of such resources must be included, varying a great deal in the range of their characteristics, but with some of them being similar to others.

[22] Thus, more of a good is preferred to less of a good when the quantity of other goods does not change; and if a basket of goods A is preferred to a basket of goods B, which is in turn preferred to a basket of goods C, then basket A is also preferred to basket C.

[23] Indeed, the presence of an income or budget constraint reflects a fundamental principle of economics, scarcity of resources.

[24] Such changes are essential for the travel-cost method. Thus, if the disappearance of grizzly bears from a park affects the sightseeing trips of visitors to the park, this disappearance is the basis for inferring that the visitors received well-being from seeing the bears. Concern about the survival of grizzlies could also be present for someone who does not visit the park, whether the grizzlies are present or not.

[25] Opaluch and Grigalunas (1991) discuss these issues. In addition, they explain how, if ethical views affect CV answers, WTP and WTA can diverge.

[26] Rappaport et al. (1981) discuss the impact of the public's misinformation on its reaction to environmental damage, as well as the media's role in the situation.

[27] For more discussion of the role of information, see the paper by Milgrom in this volume.

[28] For presentations and discussions of the CV method, see Cummings, Brookshire, and Schulze (1986) and Mitchell and Carson (1989).

[29] This change might be an improvement in the condition or a protection against the deterioration of a resource. In some surveys, respondents have been asked for their willingness to accept a deterioration in the condition of a resource. That is, instead of being asked how much they would pay to "purchase" an improvement, they are asked how much they would have to be paid to agree to the occurrence of a deterioration.

[30] The follow-up questions include asking if the information was understood and if the stated WTP was intended to make just the "purchase" of the environmental change that was described. When respondents state a zero WTP, they are asked why the answer given was zero.

[31] For a brief mathematical presentation of compensating variation measurement, see the appendix in the paper by Diamond et al. in this volume.

[32] For a discussion of choice with limited information, see the paper by Milgrom in this volume.

[33] See Urban, Katz, Hatch, and Silk (1983).

[34] See Hartman, Doane, and Woo (1991).

[35] For a discussion of the inconsistency of answers to typical yes-no questions with the economic theory of preferences, see the paper by McFadden and Leonard in this volume.

[36] In the CV study of the *Nestucca* oil spill (Rowe et al., 1991), respondents were asked the question, "Would you say the dollar amount you stated your household would be willing to pay is [followed by several alternatives]?" Only a minority selected "just for the described oil spill,"with the percentage ranging from 12.3% to 44.3% in the six reported surveys. In addition, in this survey, over a quarter of respondents left the WTP question blank or responded "Don't know."

[37] For a discussion of the generally recognized distinction between attitudes and behavior, see Ajzen and Fishbein (1977). As an example of the importance of actual payment in the findings of surveys, there is the work of Binswanger (1980; 1981). He contrasted experiments with farmers in India with the more familiar experiments usually done with American undergraduates. Because income levels in India were so much lower, he was able to offer substantial sums relative to their standard of living, an opportunity he did not have with American undergraduates. He found that the presence of more significant financial rewards had important effects on the results.

[38] See Mitchell and Carson (1989).

[39] In the more usual survey process, information can be gathered after the new product introduction or after the person votes in an election in order to ascertain if the original survey corresponded to the individual's actual decision process. With the contingent valuation of nonuse value, no actual decision process ever occurs that could provide a real-world benchmark to assess possible survey bias.

[40] Thus, the situation is analogous to specification tests in econometrics where internal consistency of a model can often be checked. However, no guarantee exists that a model that passes specification tests is reasonable without regard to real-world benchmarks.

[41] Familiarity with voting on referenda suggests that individuals will be more familiar with yes-no answers (dichotomous choice) to stated amounts to be paid than they are with stating the amount they would be willing to pay. Similarly, most purchase decisions today are yes-no decisions rather than negotiations for a bargained price. This is an issue of survey design rather than an issue of theoretical relevance.

EXHIBIT J

[42] See, for example, Magleby (1984).

[43] See the paper in this volume by Schkade and Payne.

[44] A "protest zero" is a WTP answer of zero dollars given because a respondent wishes to make a protest against the payment vehicle or some other aspect of the survey, not because the respondent truly places zero value on the good being valued.

[45] See Dickie, Fisher, and Gerking (1987). In this experiment, interviewers went to a random sample of households in Laramie, Wyoming and asked household members to state how many pints of fresh strawberries they would purchase at a proposed price for real or hypothetical transactions. Demand curves were estimated for actual and hypothetical transactions. The demand curves were very different. An adjustment was made by eliminating one set ofinterviewers and one outlier hypothetical demand. After this adjustment, demand curves were close. Because strawberries are a frequently purchased item, the relevance of the findings for nonuse valuation is limited. It is interesting that both the hypothetical and actual demand curves passed the basic test of consistency: an increase in price should decrease the quantity demanded. The coefficients on price in the estimated demand curves were statistically significant.

[46] See Bishop and Heberlein (1979). This experiment compared hypothetical and actual WTP to purchase a permit to hunt geese. Although this is not a good normally sold in this setting, it is similar to purchases of access to recreation facilities, a commonly purchased good.

[47] See Mitchell and Carson (1989).

[48] In a follow-up telephone survey of 25 of those who did not pay, 13 said that they get too many requests to support good causes and could not support everything.

[49] The study includes a third sample differing from these two in that greater efforts were made to obtain responses.

[50] We note here that the average of the hypothetical answers that were positive was roughly the same as the average of positive donations. We find this statistic supportive of the hypothesis, discussed below, that some people think in terms of charitable donations when attempting to answer CV questions.

[51] American Association of Fund-Raising Counsel (AAFRC), *Giving USA* (1991).

[52] See note 51.

[53] Eubanks and Wyckoff (1989) find that 8.5% of Minnesota taxpayers contribute to the nongame wildlife fund. Applegate (1984) finds that 5% of New Jersey taxpayers contribute, whereas Brown, Connelly, and Decker (1986) find that 5.3% of New York taxpayers contribute.

[54] Eubanks and Wyckoff (1989) find that individuals receiving refunds are more likely to contribute and contribute more when they give. Manfredo and Haight (1986) discuss Oregon's tax check-off, which is only available to those receiving refunds.

[55] See the discussion of experiments to examine free-riding in the survey by Cummings and Harrison (1992).

[56] See, for example, Mitchell and Carson (1989:168–170).

[57] This argument is presented formally in the paper by Diamond et al. in this volume.

[58] Note that if you are willing to pay all your income to protect your life, WTP increases dollar for dollar with your income. CV studies generally find that stated WTP increases less than in proportion to income.

[59] In both studies, pictures were shown to respondents in order to indicate the implications of different levels of air visibility at the Grand Canyon. For a summary of these findings, see Mitchell and Carson, (1989:46–47).

[60] For a discussion of income effects see the paper by Diamond et al. in this volume. The income elasticity of charitable donations (the percentage by which donations increase with a 1% increase in income) is plausibly similar to that of WTP amounts. Estimates of the former are typically less than one. See the survey in Clotfelter (1985).

[61] This experiment is described in detail in Diamond et al. in this volume.

[62] By asking a single WTP question, potential problems associated with the sequence of WTP questions affecting WTP answers were avoided.

[63] See Cummings, Brookshire, and Schulze (1986).

[64] See the paper by Desvousges et al. in this volume.

[65] This paper also compared open-ended and dichotomous-choice CV formats.

[66] For discussion of Kahneman and Knetsch (1992a), see Harrison (1992); Smith (1992); and Kahneman and Knetsch (1991; 1992b).

[67] This statement assumes that the allocation of their contribution to C between A and B is under their control. If the allocation between the two causes is not under their control, the inefficiency from the perspective of the individuals in the division of fund C between the causes A and B may lead to a decline in the aggregate amount contributed to fund C.

[68] Note that the issue here is again whether the answers are plausibly a reflection of preferences. This question is different from the question that has received some attention in the literature (e.g., Hoehn and Randall, 1987). In that literature, the stated WTP amounts are assumed to be a reflection of preferences, and the question asked is how the answers should be used for public-policy questions.

[69] See Kemp and Maxwell (this volume).

[70] For instance, see Balson, Hausman, and Hulse (1991); Chestnut and Rowe (1990); Rubin, Helfand, and Loomis 1991; and Desvouges et al. in this volume.

[71] For economists studying the amount given, it may not matter whether people receive satisfaction from giving or are avoiding the dissatisfaction that would come from not fulfilling an obligation that they feel.

[72] This result was found by using an open-ended CV question format. However, when we used the referendum format, no statistically significant difference was found. See McFadden and Leonard in this volume.

[73] American Association of Fund-Raising Counsel, *Giving USA, New York:* AAFRC Trust For Philanthropy (1991).

[74] They found that of a sample of 920, 34% said they had given but were not listed as donors in the Community Chest files, 31% reported not giving, 25% correctly reported giving, 8% reported giving and might have given, and 2% did not answer.

[75] See Morgan, Dye, and Hybels (1977).

[76] For example, see Keating, Pitts, and Appel (1981) and Long (1976).

[77] For modeling of charitable donations, see Andreoni (1989; 1990); Kingma (1989); and Margolis (1982).

REFERENCES

Ajzen, Icek, and Martin Fishbein, 1977, "Attitude-Behavior Relations: A Theoretical Analysis and Review of Empirical Research," *Psychological Bulletin*, 84, 5, 888–918.

American Association for Fund-Raising Counsel (AAFRC), 1991, *Giving USA*, New York: AAFRC Trust for Philanthropy.

Andreoni, James, 1989, "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence," *Journal of Political Economy*, 97, 1447–1458.

Andreoni, James, 1990, "Impure Altruism and Donations to Public Goods: A Theory of Warm Glow Giving," *The Economic Journal*, 100, 464–477.

Applegate, James E., 1984, "Nongame Tax Check-Off Programs: A Survey of New Jersey Residents Following the First Year of Contribution", *Wildlife Society Bulletin*, 12, 122–128.

Arrow, Kenneth J., and Anthony C. Fisher, 1974, "Environmental Preservation, Uncertainty, and Irreversibility," *Quarterly Journal of Economics*, May, 88, 312–319.

Balson, William E., Jeff Hausman, and Annette E. Hulse, 1991, "Navajo Generating Station BART Analysis," Prepared for the Salt River Project by Decision Focus Incorporated, Los Altos, California, April.

Binswanger, Hans P., 1980, "Attitudes Toward Risk: Experimental Measurement in Rural India," *American Journal of Agricultural Economics*, 62, 3, 395–407.

Binswanger, Hans P., 1981, "Attitudes Toward Risk: Theoretical Implications of an Experiment in Rural India," *The Economic Journal*, 91, 867–890.

Bishop, R.C., and T.A. Heberlein, 1979, "Measuring Values of Extramarket Goods: Are Indirect Measures Biased?" *American Journal of Agricultural Economics*, 61, 926–930.

Brown, Tommy L., Nancy A. Connelly, and Daniel J. Decker, 1986, "First-Year Results of New York's 'Return a Gift to Wildlife' Tax Checkoff," *Wildlife Society Bulletin*, 14, 115–120.

Chestnut, Lauraine G., and Robert D. Rowe, 1990, "Preservation Values for Visibility Protection at the National Parks," Draft Final Report prepared for the U.S. Environmental Protection Agency Office of Air Quality and Standards.

Clotfelter, Charles T., 1985, Federal Tax Policy and Charitable Giving, Chicago: Chicago University Press (for the National Bureau of Economic Research).

Cummings, Ronald G., David S. Brookshire, and William D. Schulze, eds., 1986, *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method*, Savage, Maryland: Rowman & Littlefield.

Cummings, Ronald G., and Glenn W. Harrison, 1992, "Identifying and Measuring Non-Use Values for Natural and Environmental Resources: A Critical Review of the State of the Art," unpublished.

Dickie, Mark, Ann Fisher, and Shelby Gerking, 1987, "Market Transactions and Hypothetical Demand Data: A Comparative Study," *Journal of the American Statistical Association*, 82, 397, 69–75.

Duffield, John W., and David A. Patterson, 1992, "Field Testing Existence Values: An Instream Flow Trust Fund for Montana Rivers," University of Montana, unpublished.

Eubanks, L., and J. Wyckoff, 1989, "Voluntary Contributions to State Nongame Wildlife Programs," *Journal of Environmental Economics and Management*, 16, 38–44.

Freeman, A.M., 1984, "The Quasi-Option Value of Irreversible Development," *Journal of Environmental Economics and Management*, 11, 292–295.

Hanemann, W. Michael, 1991, "Willingness to Pay and Willingness to Accept: How Much Can They Differ?," *American Economic Review*, 81, 635–647.

Harrison, Glenn W., 1992, "Valuing Public Goods with the Contingent Valuation Method: A Critique of Kahneman and Knetsch," forthcoming in *Journal of Environmental Economics and Management*, 21.

Hartman, Raymond S., Michael J. Doane, and Chi-Keung Woo, 1991, "Consumer Rationality and the Status Quo," *Quarterly Journal of Economics*, 56, 141–162.

Hoehn, John P., and Alan Randall, 1987, "A Satisfactory Benefit-Cost Indicator from Contingent Valuation," *Journal of Environmental Economics and Management*, 14, 226–247.

Kahneman, Daniel, 1986, "Comments on the Contingent Valuation Method," *in* Ronald G. Cummings, David S. Brookshire, and William D. Schulze, eds., *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method*, Savage, Maryland: Rowman & Littlefield.

EXHIBIT J

Kahneman, Daniel, and Jack L. Knetsch, 1991, "Issues of Method in the Study of Embedding: Reply to Harrison," unpublished.

Kahneman, Daniel, and Jack L. Knetsch, 1992a, "Valuing Public Goods: The Purchase of Moral Satisfaction," *Journal of Environmental Economics and Management*, 22, 57–70.

Kahneman, Daniel, and Jack L. Knetsch, 1992b, "Contingent Valuation and the Value of Public Goods: A Reply," *Journal of Environmental Economics and Management*, 22, 90–94.

Keating, Barry, Robert Pitts, and David Appel, 1981, "United Way Contributions: Coercion, Charity, or Economic Self-Interest?" *Southern Economic Journal*, 47, 815–823.

Kingma, Bruce R., 1989, "An Accurate Measurement of the Crowd-out Effect, Income Effect, and Price Effect for Charitable Contributions," *Journal of Political Economy*, 97, 5, 1197–1207.

Krutilla, John V., 1967, "Conservation Reconsidered," *American Economic Review*, 57, 787–796.

Long, Stephen H., 1976, "Social Pressure and Contributions to Health Charities," *Public Choice*, Winter, 55–66.

Magleby, David B., 1984, *Direct Legislation: Voting on Ballot Propositions in the United States*, Baltimore and London: The Johns Hopkins University Press.

Manfredo, Michael J., and Bill Haight, 1986, "Oregon's Nongame Tax Checkoff: A Comparison of Donors and Nondonors," *Wildlife Society Bulletin*, 14, 121–126.

Margolis, Howard, 1982, *Selfishness, Altruism, and Rationality*, Cambridge: Cambridge University Press.

Mitchell Robert C., and Richard T. Carson, 1989, *Using Surveys to Value Public Goods*, Washington, D. C.: Resources for the Future.

Morgan, James N., Richard F. Dye, and Judith H. Hybels, 1977, "Results from Two National Surveys of Philanthropic Activity," Research Papers, The Commission on Private Philanthropy, Department of the Treasury.

Opaluch, James J., and Thomas A. Grigalunas, 1991, "Ethical Values and Personal Preferences as Determinants of Nonuse Values: Implications for Natural Resource Damage Assessments," Peacedale, RI: Economic Analysis Inc.

Parry, Hugh J., and Helen M. Crossley, 1950, "Validity of Responses to Survey Questions," *Public Opinion Quarterly*, 14, 61–80.

Randall, Alan, and John R. Stoll, 1983, "Existence Value in a Total Valuation Framework," *in* Robert D. Rowe and Lauraine G. Chestnut, eds., *Managing Air Quality and Scenic Resources at National Parks and Wilderness Areas*, Boulder CO: Westview Press, 265–274.

Rappaport, Allen, Louis H. Zincone, Jr., and Peth Fricke, "The Media and Oil Spills: Does the Press Influence Damage Perceptions?" 1981, *Proceedings of the 1981 Oil Spill Conference, American Petroleum Institute*, 707–712.

Rowe, Robert D., William D. Schulze, W. Douglass Shaw, David Schenk, and Lauraine G. Chestnut, "Contingent Valuation of Natural Resource Damage Due to the Nestucca Oil Spill," Final Report, 1991.

Rubin, Jonathan, Gloria Helfand, and John Loomis, 1991, "A Benefit-Cost Analysis of the Northern Spotted Owl," *Journal of Forestry*, December, 25–30.

Samples, Karl C., and James R. Hollyer, 1990, "Contingent Valuation of Wildlife Resources in the Presence of Substitutes and Complements," *in Economic Valuation of Natural Resources: Issues, Theory and Applications*, ed. Rebecca L. Johnson and Gary V. Johnson, Boulder CO: Westview Press, 177–192.

Samuelson, Paul A., 1954, "The Pure Theory of Public Expenditure," *Review of Economics and Statistics*, 36, 4, 387–389.

Samuelson, Paul A., 1947, *Foundations of Economic Analysis*, Cambridge: Harvard University Press.

Schmalensee, Richard, 1972, "Option Demand and Consumer's Surplus: Valuing Price Changes Under Uncertainty," *American Economic Review*, 62, December, 813–824.

Schulze, William D., Ronald G. Cummings, David S. Brookshire, Mark A. Thayer, R. Whitworth, and M. Rahmatian, 1983, "Methods Development in Measuring Benefits of

Environmental Commodities," vol. 2, draft manuscript of a report to the Office of Policy Analysis and Resource Management, U.S. Environmental Protection Agency.

Seip, Kalle, and Jon Strand, 1991, "Willingness to Pay For Environmental Goods in Norway: A Contingent Valuation Study With Real Payment," SAF Center for Applied Research, Department of Economics, University of Oslo, unpublished.

Sen, A., 1977, "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory," *Philosophy and Public Affairs*, 6, 4, 317–344.

Shavell, Steven, 1987, *Economic Analysis of Accident Law*, Cambridge: Harvard University Press.

Smith, V. Kerry, 1992, "Arbitrary Values, Good Causes, and Premature Verdicts: A Reaction to Kahneman and Knetsch," *Journal of Environmental Economics and Management*, 22, 71–89.

Tolley, George S., Alan Randall, G. Blomquist, R. Fabian, G. Fishelson, A. Frankel, J. Hoehn, R. Krumm, and E. Mensah, 1983, "Establishing and Valuing the Effects of Improved Visibility in the Eastern United States," interim report to the U.S. Environmental Protection Agency.

U. S. Department of the Interior, Office of the Secretary, 43 CFR Part 11, 1986

Urban, Glen L., Gerald M. Katz, Thomas E. Hatch, and Alvin J. Silk, 1983, "The ASSESSOR Pre-Test Market Evaluation System", *Interfaces*, 13, 38–59.

Weisbrod, B.A., 1964, "Collective Consumption Services of Individual Consumption Goods," *Quarterly Journal of Economics*, 78, 3, 471–477.

EXHIBIT J